图书情报工作
第66卷 第11期 2022年6月

融合BP神经网络的学术话语权评价方法探讨*

■ 赵蓉英[1,2] 朱伟杰[1,2] 张兆阳[1,2] 李新来[1,2]
[1] 武汉大学中国科学评价研究中心 武汉 430072 [2] 武汉大学信息管理学院 武汉 430072

摘 要:[目的/意义]学术话语权是中国国际话语权体系中不可分割的一部分,是国家政治、经济、科学技术软实力的主要表现,对学术话语权评价方法进行剖析,综合比较不同方法的优劣与稳定性,有助于为学术话语权评价提供一定参考。[方法/过程]采用6种无需综合评价值的赋权法进行单一模型评价,对通过非参数检验后的结果进行模糊组合评价,减少单一评价倾向,提高评价公信力,并引入BP神经网络,基于梯度下降算法构建神经网络模型。[结果/结论]构建基于创新引领指数的学术引领力、基于引文分析指标的学术影响力与基于Altmetrics指标的学术传播力三维度的学术话语权评价体系,模糊Borda评价可以综合考虑单一模型的评价值大小与评价序相对秩次,实现客观信息的内部组合,相较于单一评价模型有更高的准确度,并基于此构建了融合BP神经网络的学术话语权评价模型。

关键词:学术话语权 BP神经网络 评价方法 客观赋权 模糊组合评价
分类号:G322 G250
DOI:10.13266/j.issn.0252-3116.2022.11.006

话语权可直观解释为说话的权力,表现为主体通过信息渠道等载体向受众传播信息时的现实影响力,影响着社会的发展与民众的判断。学术话语权则是在话语权基础上对其内涵进行缩小,特指学术方面的表达权,指学术流派、学术团体或学术组织的观点在学界的影响力大小或辐射范围的宽广度。因此,学术话语权本质上是指学术界所特有的术语、概念、范畴和话语体系,是对学术观点进行一定的言说和表达,由一定学术实力、学术参与度和学术规则掌握度等综合因素组成,可通过学术共同体、学术期刊以及学术数据库等媒介体现。学术话语权的形成离不开话语体系支撑,学术话语体系是由学术平台、学术质量和学术评价三者共同组合,话语体系作为传播文化和精神的支撑,需要对三大要素进行合理整合,结合价值观念和理论基础,将话语符号有序架构,形成带有自身特色的学术话语体系。

在全球化日益发展的今天,大国之间的博弈愈发激烈,话语权更是成为各国的必争之地,对学术话语权做出科学评价也是新时代中国特色评价理论与实践的要求。当前国际形势日益严峻复杂,在话语权理论基础上构建学术话语权评价模型与指标体系成为当下重要而紧迫的命题。正确使用评价方法,增强评价合理性显得尤为重要。

1 研究综述

学术话语权是国家科技创新、发展及软实力的重要组成部分,是一个国家在国际上学术话语权力结构的地位和影响力的重要体现。当前学界对于学术话语权评价的研究涉猎较少,研究大多集中在学术话语权的内涵、表现形式、存在问题等几个方面。王旭指出学界关于学术期刊话语权的现有研究"应然"呼吁多,"实然"分析少,缺乏多元融合的学术期刊话语权评价理论与实证相结合的研究,并提出从多维度、多要素、多指标、多方法融合,评价学术期刊话语权的设想[1];邵娅芬以学术话语传播者-期刊编委和话语平台-SSCI源刊为切入点,从国家和高校两个层面指出学术期刊编委、高校科研成果与一国在国际上的声誉呈显著相关关系[2];学者王兴从期刊编委角度出发,以化学学科SCI期刊为实证对象从操作层面定量对比与国外学科的差距,衡量学术话语力量的具体分布情况[3]。以上针对不同对象的多维评价分析已较为完备,但评价过程中方法选择与解释说明还较为模糊,也尚未建

* 本文系国家社会科学基金重大项目"构建中国话语权的评价科学理论、方法与应用体系研究"(项目编号:18ZDA325)研究成果之一。
作者简介:赵蓉英,教授,博士,博士生导师;朱伟杰,博士研究生,通信作者,E-mail:zhuwjzzz@163.com;张兆阳,博士研究生;李新来,博士研究生。
收稿日期:2021-11-12 修回日期:2022-04-10 本文起止页码:50-58 本文责任编辑:徐健

立起较为完善的评价指标体系。

由于学术话语权评价较为复杂，并非简单线性加权，单一评价指标难以全面涵盖信息，受限于单一视角易造成评价结果有失偏颇。也正因此，余波等人借助 Python 等分析工具处理数据，综合运用相关分析、主成分分析以及回归分析等方法，对学术话语权评价指标进行了定量分析和选取，构建了中国学术话语权评价指标体系[4]；吴坚等则将评价指标体系划分为学术生产力、影响力和引领力三方面，从多学科角度分析合作网络结构变化对体育学术话语权的影响[5]；此外，还有学者针对评价方法提出创新，如闫奕文等结合政务微信公众号信息传播特点，从用户认知、情感和态度、用户行为、社会影响和政务微信公众号平台 5 个维度，基于 BP 神经网络提出政务微信公众号信息传播效果评价方法[6]；又如俞立平提出 BP 神经网络辅助非线性评价的方法，并指出结合 BP 人工神经网络建模只能作为一种辅助手段，最终筛选仍然要结合评价目的、评价逻辑和指标的实际重要性等因素进行综合判断等[7]。

综上，尽管学者们对话语权从不同的层面、不同的角度出发进行了全面和丰富的解读，但将学术话语权与学术期刊相结合及对评价方法进行科学性检验的研究还比较缺乏，且当前多为定性研究，仅存在少量实证研究。学界针对学术话语权三大特性展开的评价具有一定的参考价值，但不能完全适用，现有评价方法多为直接建立评价体系的数学模型，采用 AHP、DEA、EWM 等，难以排除各种随机性和主观性，易造成评价结果失真和偏差。而话语权的评价过程受到众多因素影响，已有的单一评价方法不能完全应用于学术话语权评价。鉴于此，本文以 359 本经济学期刊作为实证对象，在前人研究基础上提出学术话语权的评价指标，对各项评价指标进行综合选取，遵循指标选取原则的前提下构建学术话语权的评价指标体系，在单一评价方法的基础上引入隶属度和模糊频率，运用模糊 Borda 法进行组合评价，随后引入 BP 神经网络模型，构建基于 BP 神经网络的学术话语权评价模型，以期为开展评价提供新思路和新方法。

2　相关概念与理论

2.1　学术话语权

学术话语权，是指在学术交流中拥有学术成果或学术资源的学术主体，通过各种话语表达形式对外传播其学术成果所产生的足以影响乃至引领其他行为个体认识和行动的能力[9]。学术话语权与学术组织之间是共生共存的，两者间的关系主要表现在学术思想产生影响后，不可避免地会被认定为隶属某个学术组织，反之，若是某种观点无学派支撑，相应地难以形成影响力；其次，学术话语权的产生与发展过程并不容易，需要克服其他学派思想的质疑、对比，需要经受起其他学术思想的考验、批判，被接受后方形成话语权。社会学家郑杭生对于学术话语权的看法是权力与权利的统一、资格与权威的统一，“权利”话语有多种形式，包括不受干预的学术自主研究权、供给消费等基础性意义的赋予、对物品的再造创新等权利，“权力”话语则包括意识行为的导向、主流行为的批判、个体行动的支配权等[8]。

纵观学术话语权产生过程，包含了话语权主体、话语内容及载体、话语传播媒介、话语权客体、话语权实施效果 5 个构成要素，对应话语权传播、引导、影响三大特性，借鉴赵蓉英等[9-10]对于话语权的定义和《中国经济学：话语权、范式转换及其他》[11]一文中对于话语权的阐述，将本文实证对象——学术期刊的学术话语权定义为：以具有影响力的学术主体，借相关研究平台的传播，引领学术研究潮流、影响学术研究方向、引导学界研究革新的能力。

2.2　相关理论方法

科学评价是一个动态、综合、集合的概念，广义层面是指评价科学化，狭义层面则是指以科学研究活动为对象进行的评价。随着大数据时代的到来，科学评价已应用在大数据研究中，主要体现在定量分析、定性分析以及定量和定性分析结合的评价方法上。引文分析法正是学界目前广泛应用的方法之一，引文分析多用于研究文献利用规律、科学评价、科学预测以及科学与社会的关系等方面，利用比较、归纳、抽象、概括等逻辑方法，揭示数量特征和内在规律，具有客观性、广泛适用性、简便易用性和功能特异性等特点[13]。

相较于传统引文分析计量指标，Altmetrics 被视为一种科研传播、交流和评价的新型方式，从本质上来讲是科学评价的内在需求与社交网络应用的深度结合，Altmetrics 带给学界一种以社交网络数据为基础的，更新、更快速、更全面的学术评价指标体系[14]。

然而无论是传统计量或是替代计量指标，落实于评价时均需考虑综合评价方法的运用，综合评价理论基础涵盖统计学、社会学、运筹学等多个学科，评价过程较为复杂，单一评价指标难以全面涵盖信息，多指标综合评价尽管可以利用一定的评价方法对评价对象赋予评价指数，但较为抽象。因此，本文在单一评价方法

基础上建立组合评价,考虑各赋权法计算时的偏向,有效发挥单一评价方法的优点。

同时,本文引入 BP 神经网络模型,人工神经网络已有七十多年历史,随着相关理论研究的不断深入与技术的提升,神经网络的结构趋于复杂多样。目前对 BP 神经网络模型应用较广的是多层前馈网络,基于一定学习规则按误差逆传播算法训练得到揭示映射关系的模型,引入 BP 神经网络评价思路在于通过工作信号和误差信号在前馈与反向之间的交流能够模拟指标之间关系,实现对学术话语权的评价。

3 融合 BP 神经网络的学术话语权评价

3.1 学术话语权评价指标体系

对于学术话语权的评价需要立足于话语权概念、特性及类型的解析,综合赵蓉英等[9-10]、权衡[11]对于话语权的阐述以及上文对于学术话语权的定义,本研究将学术话语权指标体系划分为学术引领力、学术影响力和学术传播力三个维度。

学术引领力指学术交流中引领领域研究前沿,掌握学术发展动态和走势,占据学术研究制高点,发挥学术引领、学术创新的能力,本文提出创新引领指数这一指标对学术引领力进行量化处理。

学术影响力指学术期刊在学界所产生的学术方面的影响,是评价学术期刊在一段特定时间内,其载文对学界或所在学科领域所产生影响的深度与广度。传统意义上的引文分析研究已有较长历史,研究方法与体系已趋于成熟,在进行科学评价时采用引文分析指标也得到较多认可,结果合理性较强。

学术传播力指期刊在传统媒介和网络时代新媒体语境下的传播能力,大数据时代,话语权更多体现于期刊如何正确引导用户就刊载的学术信息进行点赞、评论、转发、采纳、收录、索引等,这些多模态话语行为要素可由广域开放在线评价工具催生的 Altmetrics 指标数据(本文指各主流媒体、社交媒体、博客、论坛等讨论与传播平台的阅读次数、引用次数、新闻报道次数等)来体现[12],具体评价指标如表 1 所示:

表 1 学术期刊话语权评价指标体系及数据来源

序号	话语权维度	指标名称	数据来源
1	学术引领力	创新引领指数	Web of Science
2	学术影响力	总被引频次、影响因子、5 年影响因子、他引影响因子、即年指标、特征因子得分、文章影响力得分、文章总数、被引半衰期、JIF 平均百分位数、引用半衰期、CiteScore、SJR、SNIP、H5 指数、H5 中位数	Journal Citation Report、Google Scholar、Scopus
3	学术传播力	News mentions、Blog mentions、Policy mentions、Patent mentions、Twitter mentions、Peer review mentions、Weibo mentions、Facebook mentions、Wikipedia mentions、Google + mentions、LinkedIn mentions、Reddit mentions、Pinterest mentions、F1000 mentions、Q&A mentions、Video mentions、Syllabi mentions	Altmetric Explorer

3.2 数据处理

本文所研究期刊来源于 JCR 收录的“Economics”学科类别中的 373 种学术期刊,去除数据缺失与 Altmetric Explorer 中未检索到的期刊,剩余 359 本学术期刊为最终研究对象。

基于上文对于学术话语权的定义,将学术话语权评价指标体系划分的 3 个维度。即学术引领力指标从论文层面出发,借鉴科睿唯安和中国科学院联合发布的《2020 研究前沿》[15],获取经济学领域的研究前沿主题与国际经济学期刊在过去 5 年所发表的全部文献,共计 104 776 篇,采用词重叠度的方法计算两部分关键词的相似度[16-17];学术影响力测度采用学界认可度较高的引文指标,主要从 Journal Citation Report、Google Scholar、Scopus 获取;学术传播力指标主要通过 Altmetric Explorer 平台获取,日期设置为“2019.01.01 -2019.12.31”,得到 155 618 条经济学期刊替代计量指标数据值。

对期刊原始指标数据进行标准化处理,将数据转换为统一无量纲后采用极值归一化方法映射到[0,1]范围内,得到归一化数据矩阵,并采用 Kolmogorov-Smirnov 和 Shapiro-Wilk 对指标数据进行检验,Sig 值均小于 0.05,差异性显著,同时对指标进行相关性检验以筛选重复性高的指标并将其删去使各维度评价指标相对独立,在选择统计平均值与标准差的情况下进行双尾显著性检验。

学术引领力指标仅计算创新引领指数,不作增删;学术影响力评价指标体系删去 5 年影响因子、即年指标、特征因子得分、文章总数、SNIP、文章影响力得分与 H5 中位数 7 项相关系数大于 0.9 的指标;学术传播力指标在相关性分析中无明显交集,但数据存在较多 0 值,以 Altmetrics 指标数据的非零率衡量指标覆盖率,删去 Weibo mentions、LinkedIn mentionss、Pinterest mentions、F1000 mentions、Syllabi mentions 5 项指标覆盖率为 0 的指标,学术传播力维度剩余评价指标 12 项。

3.3 研究方法选择

当前学界对于多个评价指标的权重确定方法可分

为个体或群体基于主观偏好与基于客观原始数据信息两大类。客观赋权法中，粗糙集法、回归分析法在评价时需要给定综合评价意见或者评价值才能进行；而主观赋权基于决策者意图，其判断结果依赖于专家或个体的主观意见，受背景环境影响较大，难以从科学理论的角度解释结果的差异。

在评价过程中，应用不同的方法对同一期刊进行评价的结果往往存在差异，而每一种方法都各有其优缺点，难以细分每一种方法的优劣，仅采用一种评价方法无疑存在一定的局限性。本文在三维度评价指标赋权上未采用专家主观评价，且所采集原始数据只有各自属性值，无综合评价值，因此，三维度指标赋权中，采用无须事先获得综合评价值的 6 种客观方法：因子分析法[18]、主成分分析法[19]、熵值法[20]、变异系数法[21]、离差最大化法[22]、灰色关联分析法[23]进行单一模型评价，减少单一评价方法带来的误差，并基于上述方法建立组合评价，发挥单一评价方法的优点，使评价结果更加科学合理，提高评价公信力。

基于单一评价模型在对原始指标数据无量纲化处理后，为消除零值影响对数据统一平移处理，在确定评价指标权重后构造加权矩阵，对单一与组合评价模型的结果进行非参数检验，在检验一致的基础上分别进行权重赋值，综合比较各种排序结果的差异后引入隶属度，采用模糊组合的方法[24]进行组合评价，并融合 BP 人工神经网络算法，构建学术话语权评价模型。

3.4 评价方法结果对比分析

（1）因子分析法、主成分分析法。因子分析所计算得出的第一个成分的特征值为 5.620，第二个为 1.483，共同解释了国际经济学期刊原始评价指标数据的 78.931%。初始特征值总计代表数据分析过程中对指标进行划分之后所得的因子特征根，9 个因子在旋转前有 2 个特征值大于 1，由于已对学术话语权评价多余指标进行了删除，在此尽管只提取了前两个因子，仍认为是有效的。

在提取公因子的过程中，由于部分指标（成分）提取后并不能有力解释所有指标携带的信息，为了使因子载荷系数更为显著，对矩阵进行旋转，简化因子载荷矩阵结构，基于正交旋转法尽可能扩大每一列元素的距离，旋转后的成分矩阵将使各评价指标在各主成分上有更大载荷，因子旋转基于前 2 个公因子进行。

在因子提取过程中，降维带来的信息损失主要有两方面：一方面是将原始评价指标数据导入时所转换而成的原始因子，在因子分析过程中只提取了公因子并将其与原始因子进行对比，得出原始因子与公因子之间的共同度，原始因子所携带的其他信息被忽略；另一部分是在综合结果得分时，舍弃特征根低于 1 的因子，总被引与引用半衰期的信息损失量较大，均大于 0.3。

主成分分析基本步骤与因子分析相似，不再具体阐述步骤。对于学术影响力维度下的 9 个评价指标，主成分分析共提取出 2 个主成分（默认提取特征根大于 1 的成分），主成分 1 的方差解释率为 62.449%，主成分 2 的方差解释率为 16.482%，累积方差解释率为 78.931%。2 个主成分对应的权重依次为：62.449/78.931＝79.12%；16.482/78.931＝20.88%，信息损失量为除提取出的前两个成分的累计方差贡献率之外的信息。其余维度指标解读同上，不再复述，具体结果如表 2 所示：

表 2 各评价指标因子与主成分分析法结果

名称	因子 1	因子 2	综合得分系数	权重	主成分 1	主成分 2	综合得分系数	权重
特征根	5.616	1.488			5.620	1.483		
方差解释率	62.40%	16.54%			62.45%	16.48%		
影响因子	0.398 9	－0.146 2	0.284 7	12.47%	0.401 1	－0.119 9	0.292 3	12.70%
CiteScore	0.402 8	－0.096 4	0.298 2	13.06%	0.404 1	－0.069 7	0.305 1	13.25%
H5 指数	0.377 2	0.074 5	0.313 8	13.74%	0.375 5	0.099 7	0.317 9	13.81%
…	…	…	…	…	…	…	…	…
News mentions	0.446 8	0.185 3	0.263 8	10.32%	0.425 6	0.152 1	0.272 2	9.61%
Blog mentions	0.411 4	0.118 6	0.244 5	9.57%	0.386 5	0.154 1	0.246 2	8.69%

注：限于篇幅仅展示部分指标数据，News mentions 与 Blog mentions 分别代表新闻提及与博客提及（次数）

以因子分析与主成分分析作分析时常带有一定局限性[25]，因此对因子分析与主成分分析进行评价后解释力与代表性检验：两者目的都在于使所提取的因子（主成分）有明确意义，因子分析所提取的因子 1 涵盖 8 个指标，因子 2 覆盖 5 个评价指标；主成分分析所提取的主成分 1 具有较大载荷的有 11 个指标，主成分 2

有 5 个,L. R. Fabrigar 等[26]认为所提取公因子至少包含 4 个指标才能认为方法有效,涉及指标过少会造成要么信息不全,要么解释意义力度不够的问题。因此,上述两种方法都具有较好的解释力和代表性。

单调性检验如下:以两种方法所提取的因子(主成分)作为自变量,所得最终评价结果作为因变量进行岭回归,k 值取 0.01,岭回归本质是对最小二乘法的改良,放弃其无偏性且降低精度以获得更为符合实际的回归系数,两种方法的回归系数均为正,且回归拟合优度 R2 值较高,表明上述两种方法通过单调性检验。

(2)熵值法、变异系数法、离差最大化法、灰色关联分析法。熵值法弥补了上文主成分分析法中对两种关系变量的考虑不足,本文在分析各维度评价指标对 359 本经济学期刊贡献率的基础上,按熵值法原理计算各项指标的信息熵值、差异系数及权重,结果如表 3 所示,各项评价指标间权重大小有一定差异,其中总被引权重最高为 0.088,引用半衰期权重最低为 0.011;变异系数法在本文中指期刊的各评价指标数据值的变异程度,通过计算得到评价指标的标准差、平均值、变异系数与权重;离差最大化的思想在于通过原始指标提供的数据信息量计算权重,但这一方法在对指标进行赋权时采用统一公式,未凸显指标自身独立性,可能导致最终结果差异较大且兼容性低。鉴于此,本文在离差最大化的思想上引入独立权重,考虑指标独立性并提高赋权时的兼容度,采用此方法所得结果不同于其他赋权法往往赋予总被引或影响因子指标较大权重,此方法下引用半衰期的权重值最大,其次为被引半衰期,对此结果本文在最后计算赋权法之间的相关度以验证结果可靠性;在进行灰色关联度分析时,默认参考值为评价项的最大值,结合关联系数计算公式计算得出关联系数值,关联系数仅展示了参考与比较序列的相关性,总关联度由加权处理求平均值得到。

表 3 各评价指标熵值、信息效用值与权重系数

方法	熵值法			变异系数法		离差最大化法		灰色关联分析法
指标	信息熵值	差异系数	权重系数	变异系数	权重	偏导	权重	关联度
影响因子	0.955 7	0.044 3	0.088	0.826	0.098	1.004	0.089	0.920
被引半衰期	0.984 8	0.015 2	0.030	0.443	0.053	1.713	0.152	0.921
CiteScore	0.958 9	0.041 1	0.082	0.802	0.095	1.036	0.091	0.920
H5 指数	0.970 5	0.029 5	0.058	0.660	0.079	1.083	0.096	0.921
…	…	…	…	…	…	…	…	…
News mentions	0.765 9	0.234 1	0.057	2.346	0.040	1.164	0.038	0.941
Blog mentions	0.726	0.274	0.067	3.397	0.058	1.344	0.044	0.940

(3)模糊 Borda 组合评价。由单一评价方法结果可知,各评价方法对于评价指标的偏好各不相同,以学术影响力评价指标为例,在熵值法赋权倾向于总被引频次这一指标时,离差最大化法却偏向于被引半衰期,每种方法的最终评价值大小与评价序相对秩次不一致性较为明显,若对单一评价模型进行恰当组合,取长补短,就可最大限度地使用更多的有用信息,使得评价结果更为合理。因此,在上述 6 种方法的基础上,本文采用模糊 Borda 法进行组合评价。

模糊 Borda 法综合考虑了多个不同评价者对多个评价目标的优序关系及效用性[27],在实际处理数据中,采用 Kendall 检验法对单一评价结果进行一致性检验,从表 4 可知,Kendall 协调系数检验的渐进显著性为 0.000,小于 0.05,意味着采用因子分析等 6 种方法进行赋权所得结果具有关联性(一致性)。同时肯德尔秩相关系数为 0.914,大于 0.8,卡方值为 183.725,评价一致性程度较强。

表 4 Kendall 检验统计量

N	Kendall Wa	卡方	df	渐进显著性
6	0.914	183.725		0.000

在检验结果较为一致的情况下,利用模糊 Borda 组合评价法对 6 个评语集(赋权法)进行组合,引入隶属函数与模糊 Borda 数综合不同赋权法带来的得分与排序上的差异,基于模糊 Borda 计算公式[24]求得各个评价方法间的隶属函数值,依次计算隶属优度、模糊频数与模糊频率后,得到组合评价分值,见表 5。

建立学术话语权评价指标权重向量矩阵 A,并构建 359x6 的权重判断矩阵 R,得出 6 种客观赋权法的权重值,结合基于穷尽与排他的最大隶属度法则可知,学术影响力维度的综合评价结果倾向于"离差",学术传播力维度的综合评价结果倾向于"灰色关联分析",学术引领力维度评价指标为创新引领指数,不存在倾向问题。

模糊组合评价尽管确定了不同评价方法下的最佳

表 5　单一评价模型结果与组合评价模型结果对比

期刊	因子	主成分	熵值	变异系数	灰色关联	离差最大	模糊 Borda
Quarterly Journal of Economics	0.564 4	0.564 9	0.569 7	0.571 3	0.565 2	0.569 2	0.567 4
Journal of Economic Perspectives	0.432 5	0.432 7	0.434 1	0.433 2	0.432 2	0.432 9	0.432 9
Economic Geography	0.462 8	0.463 1	0.463 9	0.464 6	0.463 1	0.464 3	0.463 7
Brookings Papers on Economic Activity	0.389 4	0.398 4	0.411 3	0.414 9	0.395 5	0.404 5	0.402 0
Journal of Finance	0.366 6	0.367 1	0.372 4	0.374 1	0.367 3	0.370 8	0.369 6
…	…	…	…	…	…	…	…
Journal of Economic Literature	0.242 7	0.243 0	0.242 2	0.241 7	0.242 8	0.243 1	0.242 6
Review of Environmental Economics & Policy	0.171 8	0.171 7	0.170 3	0.170 5	0.172 0	0.171 3	0.171 3
Journal of Financial Economics	0.220 6	0.220 1	0.217 6	0.217 7	0.220 6	0.219 5	0.219 4
American Economic Review	0.284 4	0.284 4	0.282 2	0.281 9	0.284 5	0.283 8	0.283 6
Journal of Political Economy	0.237 1	0.237 6	0.240 3	0.240 2	0.237 2	0.238 7	0.238 4

注:各评价模型得分均归一化处理,省略期刊可参见 JCR

排序,但对 6 种客观赋权法进行整合仅为本文所作尝试,组合模型在计算最终结果时可能产生随机误差,或 6 种方法尽管呈现较为一致的结果但由于本文所使用的组合模型不当使最终评价结果产生偏差,因此需对使用模糊 Borda 方法的组合评价模型进行事后检验,以等级相关系数的高低确保组合评价结果与单一评价结果的一致,各评价模型间相关系数均大于 0.8,表明评价结果一致性较强。

根据多种赋权法以及模糊组合评价结果可知:排名靠前的期刊无论在传统引文分析指标或是学术传播力下的新媒体指标均有涉猎,在新闻媒体、社交媒体的提及次数均高于其他期刊,各项评价指标间的数值方差较小,即相比于排名靠后的学术期刊,这些期刊不仅重视拓展期刊影响力,还重视网络时代下期刊的与时俱进。

总体而言:美国作为学术强国,不仅在学术期刊被各大知名数据库收录数量上占有绝对优势,期刊所载文章也一定程度上按照美国所设置的规则运转,中国本土学术期刊的话语权处于绝对弱势,且网络时代下,美国众多期刊各方面都均衡发展,为期刊本身赢取了更多关注。

学术话语权综合评价结果表明引入隶属度和模糊 Borda 数能综合各评价方法排序与得分的差异,评价结果与当前国际普遍使用的复合影响因子指标值较为接近,说明本文所采用的国际经济学期刊综合评价模型能得出较为科学的评价结果,模型具有一定适用性。

综合 3 个维度综合评价发现,学术传播力指标在一定程度上可以弥补传统引文分析指标的不足,是对引文分析评价指标的一种扩充。网络时代下,学术期刊与网络的融合也逐渐成为新趋势,传统指标的不足推动新指标的发展,而新指标也将会为传统评价带来新的发展动力。

3.5　BP 神经网络评价模型

(1)模型建构。当前标准的 BP 神经网络已经广泛应用到各个领域,并取得了较多的研究成果,其学习算法和模型函数较为成熟,可借鉴应用于学术话语权评价。引入 BP 神经网络算法能够避免人为主观性因素的影响,而且可以通过神经网络的训练,找寻输入和输出之间的内在联系,以权重的形式保存在神经网络中,根据学术话语权不同主体类型设计个性化评价体系,增加评价模型的适应性和通用性。

本文从模型结构与权值阈值入手,考虑单隐层逼近能力与学术话语权评价的网络规模,本研究设计一层隐含层,以 3 个不同维度下筛选后的二级评价指标为输入层节点数,输出层神经元个数为评价结果的个数,而隐含层数量过拟合或欠拟合都会影响预期效果,数量的多少直接决定了网络的精度决策边界、拟合映射平滑度、网络泛化能力与训练时长,为满足精度要求且结构尽可能紧凑,采用经验公式法与试凑法确定具体数量,常用的隐含层神经元个数确定经验公式为:

$$M = \sqrt{m + n} + a \qquad \text{公式(1)}$$

其中,n 和 m 分别表示输入、输出神经元的个数,a 为 1 至 10 的常数,根据试凑法通过 4 至 8 个节点分别进行训练计算,以迭代过程中的网络均方差最大值确定隐层节点数,使用 Neural Network 工具箱构建 5 个 BP 神经网络进行建模,结合平均误差局部最小和 StackOverflow 取值范围确定隐层神经元个数为 5,具体评价模型见图 1。

由于原始数据已经过二维映射处理,因此隐含层采用传递函数 tansig,输出层采用 logsig,学习函数为默

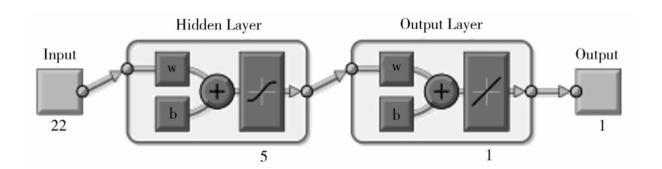


图 1 学术话语权评价模型神经网络拓扑结构

认的 learngdm,训练函数为 trainlm 函数,误差训练方法采用列文伯格 - 马夸尔特法(Levenberg-Marquardt),MSE 采用默认函数。

$$MSE(y,y') = \frac{\sum_{i=1}^{n}(y_i - y_i')^2}{n} \qquad 公式(2)$$

MSE 表示真实值与预估值差平方的期望,其值越大,预测效果越差。输入神经元数量为 22,输出层神经元数量是 1,隐层神经元数量是 5,指定初始权值和阈值将使参数与权值朝同方向修正,因此随机选取 0 至 1 间的值,BP 神经网络学术话语权评价模型训练流程如图 2 所示:

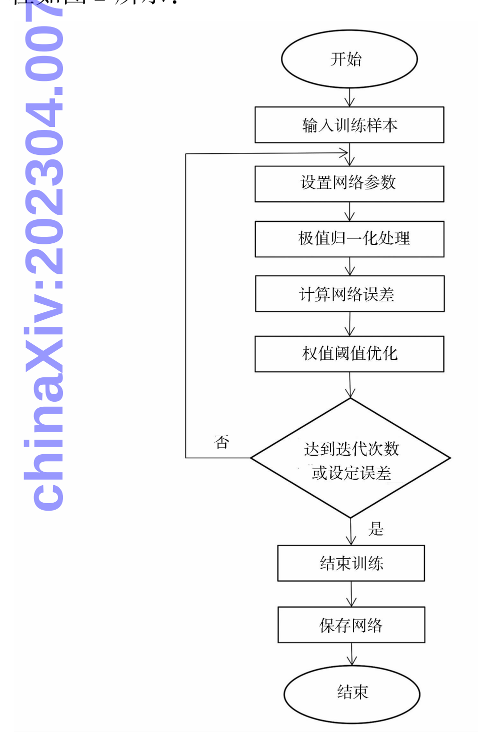


图 2 BP 神经网络学术话语权评价模型流程

在评价模型的样本设计上,按照 8:2 的比例将国际经济学期刊数据随机划分为两组数据,其中 80% 数据作为训练集,20% 数据作为测试集,考虑参数与数据集,误差足够小的理想情况较难实现,设定最大迭代次数为 500,采用 22 - 5 - 1 的单隐层 BP 网络,基于训练数据进行学习,达到目标误差后停止训练,将此时的网络权值与阈值作为学术期刊学术话语权评价的最优网络模式,并使用这一模式验证网络模型可行性,具体训练参数设置如下:

```
net.trainParam.show = 50
net.trainParam.lr = 0.01
net.trainParam.mc = 0.9
net.trainParam.epochs = 500
net.trainParam.goal = 0.000 1
[net,tr] = train(net,P,T)
```

(2)结果分析。在对神经网络初始化参数的基础上,将训练集数据输入、进行训练,将训练样本指标数据输入,调用设置好的网络训练参数进行神经网络模型的训练,由于数据集较少,训练次数较少,达到网络训练误差平方和 MSE 达到目标误差要求后模型的输出值与期望值较为接近,通过对实际输出和期待输出平均误差的计算,在达到平均误差目标后训练循环停止,由于只添加一层隐含层,训练通道未增加,误差值(默认为 0)为 0.000 1,实际过程中往往达不到最佳理想情况,且误差过小还容易出现过度拟合的情况,在训练集中,期待输出值与模型实际输出值之间的关系见图 3。

从图 3 中可以发现,本文构建的基于 BP 神经网络的学术话语权评价模型对训练集的评价效果较好,期待输出值与实际输出值之间的相关系数 R 为0.999 96。

通过以上计算,本研究共得到以下结果:

首先,基于 BP 神经网络算法构建学术话语权评价模型,通过对训练集的训练可以达到预设精度要求,BP 神经网络评价模型在学习效率与样本验证效果上满足期刊学术话语权评价,基于梯度下降优化算法的网络模型预测率达到 80% 以上,说明利用 BP 神经网络构建学术话语权评价模型是可行的。其次,基于 BP 神经网络训练集的不断自学习获得的权重能够避免人为赋予指标权重的主观性,同时减少指标之间的互相干扰,由测试集误差值可看出基于神经网络算法的期刊学术话语权评价模型具有较强实用性。最后,采用 22 - 5 - 1 的单隐层 BP 网络,在达到预设精度要求的情况下,经过训练得到各指标的权重,这在一定程度上是对所采用评价方法权重倾向的映射,基于梯度下降算法的 BP 人工神经网络与模糊组合评价存在高度相关性,表明训练好的神经网络模型可根据新输入的数据,得到评价对象的评价结果。

4 结语

本文提出融合 BP 神经网络算法的学术话语权评

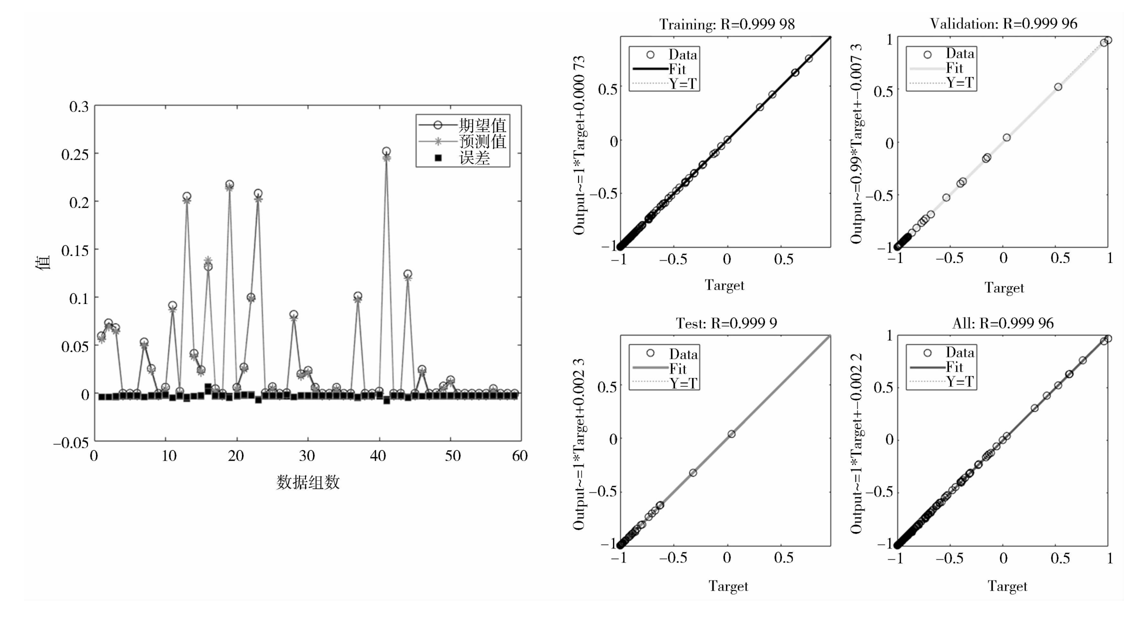


图 3　训练集期待输出值与实际输出值的对比情况与相关性分析

价模型，首先采用 6 种单一评价方法与模糊 Borda 组合评价法对经济学期刊学术话语权进行了综合评价探索，结果呈现一定差异，单一客观赋权法所得的评价结果均通过了非参数检验，表明 6 种方法具有较高一致性，在此基础上所建立的模糊组合评价模型显示其结果综合考虑了评价值大小与评价序相对秩次，平衡结果差异，可充分利用有效信息，克服单一评价方法的片面与局限，使得评价结果更为合理。引入的 BP 神经网络模型误差率达到了预设精度，期望结果与实际训练结果存在较高相关性，表明构建的学术话语权评价模型可应用于实际评价中，组合评价的方法可服务于学术话语权综合评价，适用于相关领域的评价问题，可以为研究定量评价或综合评价问题提供借鉴和参考。

6 种客观赋权方法对评价指标分别赋予了不同权重，但每一种方法均存在一定不足，总体而言，样本数据与实际问题域都将直接影响属性重要度，本文中确定权重的方法、综合评价方法、组合评价方法完全依赖于数据本身所含信息，尽管可避免人为干扰，但当评价指标之间出现信息重复的问题时，模糊综合评价需要与别的方法结合使用，且隶属函数的确定还没有系统的方法，隶属度的计算仅是对实际评价值的区间线性变换。

相较于传播效果等典型非线性评价，借助 BP 神经网络的纠错与规避误差功能以达到学术话语权评价的有效性仍需进一步探讨，其性能与训练样本直接相关，训练样本数、输入样本质量都对结果有一定影响，由于 BP 评价模型的隐含性，输出层评价结果与输入层评价指标并非简单线性关系，不能解释为回归系数，其结果还需从神经元个数、节点数、函数分布密度等角度进一步改善。

此外，本文所列举的评价指标为当前学界较为推崇的指标，在后续研究中，如何选择合适、科学、正确的期刊指标、建立完善的评价模型，仍是需要深入研究的课题。只有在研究中逐步完善评价指标体系和评价方法，避免主观获取数据，科学取样，评价才能更有说服力。

参考文献：

[1] 王旭. 中国学术期刊话语权评价理论框架建构研究[J]. 图书情报工作,2021,65(12):83－92.

[2] 邵娅芬. 经济学科的国际学术话语权研究[D]. 上海:上海交通大学,2011.

[3] 王兴. 国际学术话语权视角下的大学学科评价研究——以化学学科世界 1387 所大学为例[J]. 清华大学教育研究,2015,36(3):64－75.

[4] 余波,赵蓉英,邱均平. 中国学者学术话语权评价指标体系的构建与应用研究——以国际生物学为例[J]. 情报科学,2021,39(12):105－112,125.

[5] 吴坚,贺蕾,盘劲呈,等. 体育学学者学术合作网络对学术话语权影响的实证研究[J]. 武汉体育学院学报,2022,56(1):28－35.

[6] 闫奕文,张海涛,孙思阳,等. 基于 BP 神经网络的政务微信公众号信息传播效果评价研究[J]. 图书情报工作,2017,61(20):53－62.

[7] 俞立平. 基于神经网络的非线性学术评价方法选择研究[J]. 情报理论与实践,2021,44(1):63－70,56.

[8] 郑杭生. 学术话语权与中国社会学发展[J]. 中国社会科学,2011(2):27－34,4,220.

[9] 赵蓉英,张晓曦,刘卓著,等. 话语权与话语权评价解析[J]. 情报理论与实践,2021,44(11):15－22.

[10] 赵蓉英,王旭,余波,等. 中国话语权的评价科学研究框架构建[J]. 图书与情报,2019(4):122 – 131.
[11] 权衡. 中国经济学:话语权,范式转换及其他[J]. 探索与争鸣,2006(3):38 – 40
[12] 王旭,赵蓉英. 融合 Altmetrics 的话语引导力指标与被引频次关系研究——以中国英文学术期刊为例[J]. 情报理论与实践,2021,44(7):37 – 43,36.
[13] 赵蓉英,魏绪秋,王建品. 引文分析研究与进展[J]. 情报学进展,2018,12:50 – 80.
[14] 杨岳,王贤文. Altmetrics 十年发展综述[J]. 情报杂志,2021,40(11):136 – 146.
[15] 科睿唯安. 2020 研究前沿[EB/OL]. [2021 – 06 – 27]. https://solutions. clarivate. com. cn/blog/20201113/.
[16] 王春柳,杨永辉,邓霏,等. 文本相似度计算方法研究综述[J]. 情报科学,2019,37(3):158 – 168.
[17] 于鹏. 逻辑公式间的 Jaccard 距离及其应用[J]. 计算机科学与探索,2020,14(11):1975 – 1980.
[18] SUN J. A note on principal component analysis for multi-dimensional data[J]. Statistics & probability letter,2000(46):69 – 73.
[19] LEIGH S. A user's guide to principal components[J]. Technometrics,1993,35(1):83 – 85.
[20] CHEN M, LU D, ZHA L. The comprehensive evaluation of China's urbanization and effects on resources and environment[J]. Journal of geographical sciences,2010,20(1):17 – 30.
[21] GUPTA R, TRIPATHI R, MICHALEK J, et al. An exact test for the mean of a normal distribution with a known coefficient of variation[J]. Computational statistics & data analysis,1985, 3:219 – 226.
[22] 王应明. 运用离差最大化方法进行多指标决策与排序[J]. 系统工程与电子技术,1998(7):26 – 28,33.
[23] 谭学瑞,邓聚龙. 灰色关联分析:多因素统计分析新方法[J]. 统计研究, 1995(3):46 – 48.
[24] 徐林明,林志炳,李美娟,等. 基于模糊 Borda 法的动态组合评价方法及其应用研究[J]. 中国管理科学,2017,25(2):165 – 173.
[25] 俞立平,刘骏. 主成分分析与因子分析法适合科技评价吗?——以学术期刊评价为例[J]. 现代情报,2018,38(6):73 – 79,137.
[26] FABRIGAR L R, WEGENER D T, MACCALLUM R C, et al. Evaluating the use of exploratory factor analysis in psychological research[J]. Psychological methods,1999, 4(3):272 – 299.
[27] 熊国经,熊玲玲,陈小山. 组合评价和复合评价模型在学术期刊评价优越性的实证研究[J]. 现代情报,2017,37(1):81 – 88.

作者贡献说明:
赵蓉英:论文选题、框架设计、论文指导;
朱伟杰:数据分析、论文撰写、论文修改;
张兆阳:数据获取、数据处理;
李新来:数据分析、论文修改。

Research on the Evaluation Method of Academic Discourse Power by Fusing on BP Neural Network

Zhao Rongying[1,2] Zhu Weijie[1,2] Zhang Zhaoyang[1,2] Li Xinlai[1,2]
[1] Research Center for Chinese Science Evaluation, Wuhan University, Wuhan 430072
[2] School of Information Management, Wuhan University, Wuhan 430072

Abstract: [**Purpose/Significance**] Academic discourse power is an integral part of China's international discourse power system and the main manifestation of national political, economic, scientific and technological soft power. Analyzing the evaluation methods of academic discourse power and comprehensively comparing the advantages, disadvantages and stability of different methods are helpful to provide certain references for the evaluation of academic discourse power. [**Method/Process**] Six assignment methods without comprehensive evaluation values were used for single model evaluation, and fuzzy combination evaluation was carried out on the results after passing the non-parametric test to reduce the tendency of single evaluation and improve the credibility of evaluation, and BP neural network was introduced to construct neural network model based on gradient descent algorithm. [**Result/Conclusion**] This paper constructs an academic discourse power evaluation system based on three dimensions: academic leadership based on the innovation leading index, academic influence based on the citation analysis index and academic communication capacity based on the Altmetrics index. Fuzzy Borda evaluation allows for a combination of high and low evaluation scores and evaluation ranking for a single model, and realizes the internal combination of objective information, which has higher accuracy than the single evaluation model. Based on this, an evaluation model of academic discourse power integrated with BP neural network was constructed.

Keywords: academic discourse power BP neural network evaluation method objective empowerment fuzzy combination evaluation